ASSOCIATION BY SYNAPTIC FACILITATION IN HIGHLY DAMPED NEURAL NETS

E. M. HARTH *and* S. L. EDGAR

From the Department of Physics, Syracuse University, Syracuse, New York 13210

ABSTRACT Cognitive functions are sought in a homogeneous, randomly connected net of neuron-like elements. Information is assumed to be contained in the instantaneous states of the system, which specify the firing states (*off* or *on*) of each neuron in the net. The hypothesis of synaptic facilitation is assumed to be the basis of learning and memory. Owing to the high degree of damping no reverberations occur in the net. However, close analogies can be found between the performance of the net and known association functions of the cerebral cortex, among them various types of conditioned reflexes. The data are obtained by a combination of mathematical analysis and computer simulation. It is emphasized that the biological entity simulated by this model is at best a limited component of the cerebral cortex.

INTRODUCTION

The field which is concerned with the way the brain accomplishes some of the so-called higher functions, is a unique area of intersection of many classical disciplines, each one rich in accumulated knowledge, and each possessing its own jargon. It is possible, however, to name a few landmarks of crucial experimental data, or of particularly enlightening thought. Among these are the discovery of the main anatomical features of the cerebral cortex by Ramón y Cajal (1909) and Lorente de Nó (1949), the physiology of individual neurons (Hodgkin and Huxley, 1952) and that of synapses (Eccles, 1953, 1964), the cortical ablation experiments by Lashley (1950), Penfield's (1950) studies on epileptogenic foci and local cortical stimulation in humans, and the experiments by Burns (1958) on neuronal activity in isolated slabs of cortical material. From the behavioral sciences we have the discovery and elaboration of the conditioned reflex by Pavlov (1960) and his followers, and Hebb's (1949) many brilliant attempts at interpreting psychological functions in terms of physical models, in particular the postulate of synaptic change and his cell assembly theory. Finally, on the theoretical side, there are the analogies between computer and brain (von Neumann, 1958), the application of symbolic logic to neural functioning (McCulloch and Pitts, 1943), Rosenblatt's perceptron (1962), and Caianiello's studies of reverberatory neural nets (1961). The list is somewhat arbitrary and certainly not exhaustive.

In many quantitative studies investigators sought to establish brain-like functions by computer simulation or mathematical analysis. This work was to a large extent inspired by Hebb's ideas. According to the hypothesis of synaptic facilitation (Hebb, 1949) the effectiveness of a snyapse in triggering the postsynaptic neuron is enhanced whenever a presynaptic action potential succeeds in firing the cell. This facilitation of selected pathways was believed to constitute the physical basis of memory.

If memory is to be established in the network as a result of a single experience, it was argued that mechanisms must exist which produce sustained reexcitation of the same neural pathways following the sensory input. Hebb postulated that this reexcitation can be provided in a richly connected network through reverberations, and that by an interaction of reverberations and synaptic facilitation, a subset of the neural net would differentiate itself into what he called a *cell assembly*. This would represent something like the emergence of meaning in the firing patterns of cortical neurons. Accordingly much subsequent work was designed to establish the existence of quasi-stable modes of sustained neural activity.

Several difficulties in this attractive scheme soon appeared. Beurle (1956) and later Ashby, von Foerster, and Walker (1962) pointed out an inherent instability, according to which neural nets would tend toward one of two stable states, namely complete cessation of activity or saturated activity. The information content of the net would thus be reduced to just one bit.

Stimulation studies of neural nets by Rochester et al. (1956) produced what they termed *diffuse reverberations*, aperiodic activity involving virtually the entire network, with no tendency toward cell assembly formation. Similar results were obtained by Farley and Clark (1961). Oscillatory and quasi-stable conditions between the two extremes mentioned were obtained under certain conditions and for specific choices of network parameters (Farley and Clark, 1961; Smith and Davidson, 1962; Griffith, 1963). An excellent review of this field has recently been published by Harmon and Lewis (1966).

Model for an Association Network

In the present work we wish to retain Hebb's concept of synaptic facilitation and, though in somewhat modified form, his idea of cell assemblies. Our departure, in brief, will be as follows:

We believe that in an undifferentiated net the ability on the one hand to sustain reverberatory activity without loss of information, and the ability, on the other hand, to carry out associative functions, are complementary. Thus, the fact that cues from different sensory modalities are readily associated in the brain would seem to require a richly cross-connected net, which indeed exists in the so-called association area of the cortex. The simplest model of such a net is a randomly connected net in which the terminals of neural interconnections are chosen without geometrical

bias. If the parameters in such nets (multiplicity of connections, thresholds) are chosen to produce sustained activity following excitation of a number of cells, the work of Rochester et al. demonstrated that rapid loss of information would ensue. On the other hand, simulation studies by the same authors, in which the connectivity was given a strong local bias, did demonstrate fractionation into cell assemblies. Similar results were obtained by White (1961). Such nets, however, would be expected to make poor associators.

In view of these considerations we have decided to separate the two functions, i.e., to relieve the associator net of the requirement to sustain prolonged but limited activity. If such reverberations are necessary for the laying down of a memory trace, this may perhaps be accomplished by strongly coupled loops reaching outside the associator system, for example by the cortico-thalamocortical connections which are known to exist (Fair, 1963).

In the present paper we shall restrict our attention to the study of learning by association and the type of information processing which might take place at the highest level of cortical integration. We consider for this purpose a network which exhibits random connectivity, without bias favoring connections to nearest neighbors. We assume the property of synaptic facilitation. Network parameters are chosen in such a way that activity in the isolated associator net would almost immediately cease, following an initial excitation. In this sense the net is highly damped. It has no capacity for short-term retention of information.

In the following, and for the purposes of this paper only, we shall call this hypothetical network of neurons the *cortical net.* Activity in this net is described simply by stating which set of cortical neurons are made active as a result of a sensory input. The specification of this set shall be called a *cortical state* (CS). This set of neurons takes the place of the more dynamical concept of a cell assembly. It comes closer, perhaps, to what Zeeman (1962) has called a *thought.* We distinguish further between fundamental cortical states (FCS) and acquired cortical states (ACS). The FCS is the cortical state which results directly from the activity in the sensory net, whereas the ACS includes in addition neurons triggered by the intracortical connections.

We define a sensory state (SS) as a set of sensory neurons which are simultaneously triggered by a stimulus. The synaptic couplings that lead from the sensory to the cortical net are assumed not to be subject to changes; learning in our model is strictly confined to the cortical net, affecting only the intracortical synapses. The mapping from SS to FCS, assuming the cortical net to have been quiescent at the outset, is thus invariant, while the transformation from FCS to ACS will be a function of time. This is shown schematically in Fig. 1. The ACS therefore contains information not only about the stimulus which caused it but also about the history of the network. We shall therefore focus our attention on the ACS's, developing criteria for their comparison and description, and attempting to find in their relationships simple, readily definable psychological functions.

Outline

Some of the features of the network under study here were described in a previous paper we shall refer to as *I* (Harth 1966a). The structure of the net is essentially the same as that described in *I*, except for the addition of inhibitory connections. A description of the structure and the assumed dynamical properties of the net will be given under the Description of the Model.

Under Nomenclature and Calculations we present some calculations and preliminary computer runs designed primarily to explore the dependence of network behavior on various parameters. The curves obtained, some empirical and some calculated, help in selecting sets of network parameters for more detailed study by computer simulation, and form much of the guidelines for predicting and interpreting network behavior.

Under the Association of Cortical States we define the association of two or more cortical states and explore some of the logical structures that may be formed by

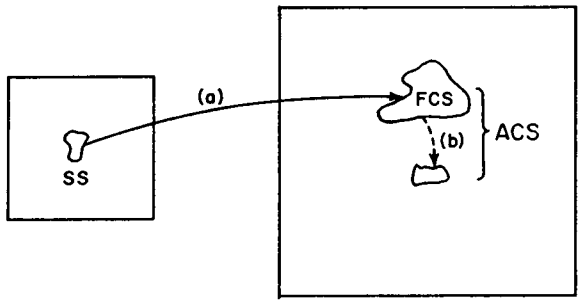

FIGURE 1. Schematic of mapping between sensory and cortical states. (*a*) invariant mapping SS → FCS; (*b*) time dependent transformation FCS → ACS.

these associations. Under the Computer Simulation of Cognitive Tasks it is shown that the system is capable of performing simple cognitive tasks including various types of conditioned reflexes. A simple association diagram, derived from the concept of the connectivity matrix, is presented as an aid in understanding the functioning of the system.

Under Limitations and Extensions we attempt to show that the present model, even though it represents only a radically simplified picture of one component of the brain, is capable of being extended, and may be fitted at a later stage into a more complete theoretical structure.

DESCRIPTION OF THE MODEL

The ingredients of the model come from many sources and have been used by other investigators in various combinations. The system to be investigated here is an abstraction; its parts are named after biological entities which they are designed to resemble in certain respects.

The system on the whole is meant to simulate that portion of the brain whose function has been the most poorly understood: the cerebral cortex—more specifically, those portions of the cortex which have been referred to, vaguely, as association areas. Whether of not the few points of similarity between our model and its physical counterpart are significant, and whether the differences (of which there are many) are trivial enough not to cancel the former, will determine how useful a model we have.

The Neuron

The model consists of a network of similar, decision-making elements, the neurons. Each neuron is in one of two possible states, off or on (the latter is also called *active*, or *firing*). The neuron consists of an afferent area comprising a cell body and dendritic tree, and an axon with axonal branches making up the efferent portion. When the neuron is on, an electrical signal of standard height, the action potential, is generated at the base of the axon, and will travel outward along the axon without attenuation. It will be assumed in this model that at axon branchings the same standard signal appears and is propagated along each branch.

Axonal branches terminate on cell bodies or dendritic branches of other neurons, forming specialized structures called *synapses*. A synapse will cause a signal, the postsynaptic potential (PSP), to be generated on the membrane of the postsynaptic neuron, whenever an action potential arrives at the presynaptic membrane. We assume the action potential to have unit height and call the size of the PSP the *coupling coefficient* k_{ij} ; here the ith neuron is the receiving, or postsynaptic neuron, the jth neuron is presynaptic. To simplify the language we consider the coupling from neuron j to neuron i to be due to a single synapse of strength k_{ij}, although this coupling coefficient may also be viewed as the sum of coefficients of any number of junctions between this pair of neurons.

The values of k_{ij} may be positive or negative, depending on the identity of the jth neuron. A neuron is called *excitatory* if all of its efferent synapses produce positive PSP's and *inhibitory* if the synapses produce negative PSP's.

All PSP's generated in the afferent area of a neuron are added linearly regardless of where on the dendritic tree or cell body they are located. If this sum exceeds a fixed value, the *firing threshold*, the neuron will go into its active state. All timings are disregarded in our treatment. This point is further discussed under Limitations and Extensions.

The Net

When a number of neurons are interconnected, the connectivity may be expressed by the matrix $\{k_{ij}\}$, which we call the *connectivity matrix*. We adopt the notation that a vanishing matrix element $k_{mn} = 0$ denotes that no axonal ending from the nth neuron synapses with the mth neuron. By a random net we mean one whose nonvanishing matrix elements are randomly distributed over the area of the matrix.

The net may be structured, consisting of distinct components such as the hypothetical net shown in Fig. 2, in which there are synapses (shown by dots) leading from sensory to association neurons (but not vice versa), between association neurons, and from association to motor neurons. In this study we shall be primarily concerned with an undifferentiated random association net (area *2* in Fig. 2) which receives inputs from a smaller sensory net (area *1* in Fig. 2).

Learning

We shall assume synaptic facilitation to be the mechanism by which memory is acquired and stored. In our model this shall mean that a positive coupling coefficient

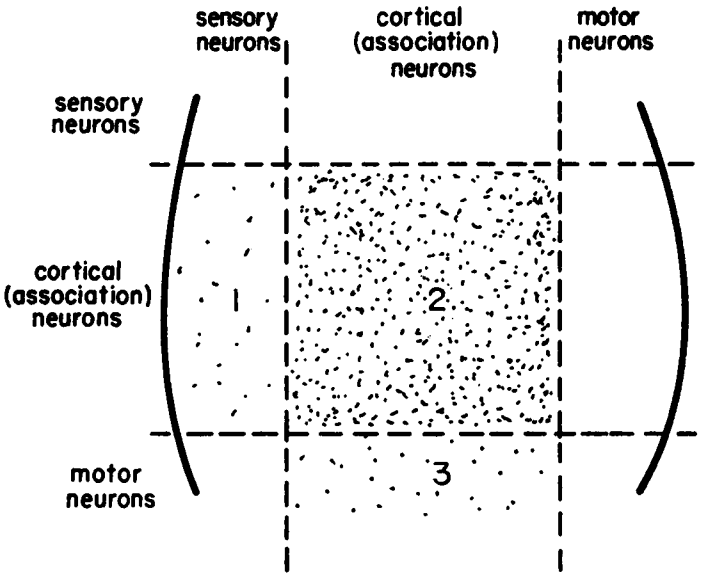

FIGURE 2. Connectivity matrix for a 3-component neural net. Dots represent the distribution of nonvanishing matrix elements.

k_{ij} will be increased by an amount δ as a result of the simultaneous firing of neurons i and j, i.e., whenever a presynaptic action potential coincides with a postsynaptic firing. This plasticity shall be restricted to intracortical synapses (area *2* in Fig. 2) and further only to those whose presynaptic neurons are excitatory. Clearly no facilitation occurs for $k_{ij} = 0$ which denotes the absence of a synaptic link.

The changes are assumed to be slow so that they may be disregarded during the development of a cortical state. This was expressed formally by Caianiello (1961) as the *adiabatic learning hypothesis*. Thus repeated occurrence of a stimulus and its resulting cortical state would be required to alter the coupling coefficients appreciably; or else a single stimulus may cause repetitive occurrence of the same or similar cortical states by continued mental preoccupation with the event. This is not the

type of simple reverberation that Hebb made responsible for the formation of his cell assemblies. We shall not be concerned here with the mechanism by which the repeated presentation of a cortical state is effected, but simply assume that such a mechanism exists. Accordingly we may reinforce a given cortical state, assuming that it has occurred a large number of times either by repetition of the stimulus or by other means.

NOMENCLATURE AND CALCULATIONS

The nomenclature used here will follow fairly closely that adopted in *I*. A cortical state (called *output state* in *I*) may be thought of as a column matrix

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_A \end{pmatrix}$$

where a_i equals zero if the ith neuron is in the resting state, and one if it is triggered; A is the total number of neurons in the net. Similarly the SS is given by

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \cdot \\ \cdot \\ s_S \end{pmatrix}$$

Here S is the total number of sensory neurons. The connectivity of the net is given by the matrix $\{k_{ij}\}$ whose elements are the coupling coefficients. We consider here (unlike *I*) only the neurons triggered after a single synaptic delay following the sensory input. This leads to considerable simplification, and, in the case of strong damping, will take care of practically all of the neurons triggered.

The thresholds t of all neurons will be assumed to be identical. If the combined input $\sum_{j=1}^{A} k_{ij}a_j$ for the ith neuron exceeds t it will go into the firing state $a_i = 1$.

The values of the nonvanishing matrix elements are changed according to the reinforcement rules which can be stated

$$k'_{ij} = k_{ij} + \delta_{ij}\,; \qquad \delta_{ij} = \begin{cases} \delta \text{ iff } a_i = a_j = 1; \quad k_{ij} > 0 \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

Clearly the k matrix represents at any moment the connectivity as well as the past history of the system while the present activity is expressed by the CS.

The relationships between the *SS* and resulting cortical states was explored in I and in subsequent experiments whose results will only be outlined here. In the present paper we are more directly concerned with the cortical states.

It was already pointed out that under our assumptions the FCS is an invariant result of a given SS. The FCS's shall be denoted by the symbols a_0, b_0, etc. and their corresponding time dependent ACS's by a, b, etc. It is assumed that triggering of the additional neurons is instantaneous, so that the state **a** always contains all neurons active in a_0. We write $a_0 \subset a$, $b_0 \subset b$, etc.

By the size of a state we mean the number of active neurons contained in it; we denote by S the size of an SS and by $_a$ the size of a CS. Thus

$$S = \sum_{i=1}^{S} s_i$$

$$_a = \sum_{i=1}^{A} a_i$$

We also introduce the specific sizes σ and α defined by $\sigma = S/S$, $\alpha = {}_a/A$. Curves of α vs. σ were given in I for a net having only excitatory connections.

Network Parameters

The nets to be studied are generated by the computer from a set of selected parameters. These are:

- S the number of sensory neurons
- A the number of cortical neurons
- $\bar{m}$ the average number of afferent synapses per cortical neuron, originating in sensory net
- $\bar{\mu}_+$ the average number of intracortical efferent synapses per excitatory cortical neuron
- $\bar{\mu}_-$ the average number of intracortical efferent synapses per inhibitory cortical neuron
- h the fraction of inhibitory neurons in the cortex
- c the coupling coefficient for sensory-cortical synapses
- k_+ the initial coupling coefficient for excitatory intracortical synapses
- k_- the coupling coefficient for inhibitory intracortical synapses
- t the firing threshold of all neurons

In addition the distribution of the values for m, μ_+ and μ_- about their respective averages must be given. Having chosen a set of parameters, the computer generates the net by taking each cortical neuron in turn and distributing the origins of its m sensory-cortical connections randomly among the S sensory neurons, and the addresses of its μ_+ (or μ_-) efferent synapses among the A cortical neurons.

Comparison of Cortical States

As in I, we are not concerned here with motor outputs which ultimately result from neuronal activity, but rather with the states of those portions of the neural net, our FCS's and ACS's which precede the strong convergence of neuronal activity. The system we are simulating is thus not a decision-making device like the perceptron (Rosenblatt, 1962).

The ACS's which must carry considerably more information than the motor outputs, are, however, not as readily interpreted as the latter; one reason for this is the fact that in the real brain they are not subject to direct observation. We must therefore make up our own interpretation concerning the meaning of these states, and develop a method for comparing them with one another.

Viewed as subsets of the total set of cortical neurons, two states a and b may have one and only one of the following relations:

(1). $a \subset b, \quad a \neq b$ (2). $a \supset b, \quad a \neq b$ (3). $a = b$ (4). $a \not\subset b, \quad a \not\supset b.$

It is reasonable to suppose that relation 4 is the most frequent between states. Two such states are called *incomparable* elements in set theory (Birkhoff and MacLane, 1953). It will be out task, nevertheless, to compare such states. We propose to do this by defining as *related* those states which have a nonvanishing Boolean overlap $a \cap b$. The relation is stronger the greater this overlap. Clearly any subset of a state a is related to a. This relationship may be likened to the equivalence between a hologram and any of its subsets (Leith and Upatnieks, 1964). Relatedness as defined here is clearly a binary relation obeying the reflexive and symmetric laws; it shall be expressed by a $\mathsf{R}\ b$ (a is related to b). States whose overlap is the null set are unrelated or disjoint states.

It will also be convenient to have three correspondences between elements of the class of sensory states on the one hand and cortical states on the other. The first is a one-one correspondence called *meaning*. A cortical state *means* s if it is the FCS resulting from s. We write $a_o\ \mathsf{M}\ s$, a_o means s.

We shall say that a cortical state b *implies* s, in symbols $b\ \mathsf{I}\ s$, if there exists a state c such that $c \subset b$ and $c\ \mathsf{M}\ s$.

Finally, a state b *suggests* s, in symbols $b\ \mathsf{S}\ s$, if there exists a state c such that $b\ \mathsf{R}\ c$, $c\ \mathsf{M}\ s$, and $(b \cap c)$ is a proper subset of c.

It follows that an arbitrary cortical state may have at most one meaning, but imply and/or suggest more than one sensory state.

Calculations of Net Behavior

We have found a combination of mathematical analysis and computer simulation fruitful in exploring the behavior of simple nets. Under the assumptions made we can readily calculate the expectation values $\langle \alpha_o \rangle$, the fractional number of cortical

neurons belonging to the FCS which results from an SS of specific size σ. The averages are over possible nets and possible inputs. One finds that with the choice of parameters, $\langle \alpha_o \rangle$ does not explicitly depend on the net sizes S and A, nor even on the ratio of the two. The relationship is particularly simple if all sensory-cortical connections have the same coupling strength. Since the average number of active sensory-cortical connections per cortical neuron is $(\sigma \bar{m})$, we obtain from the Poisson formula

$$\langle \alpha_o \rangle \approx 1 - \exp(-\sigma \bar{m}) \sum_{l=0}^{\xi-1} (\sigma \bar{m})^l / l! \tag{2}$$

Here ξ is an integer giving the minimum number of inputs required to trigger a neuron. We introduce here the function $u(x)$ which is the largest integer equal to or smaller than x. Thus $\xi = u(t/c) = t/c - t/c \pmod 1$.

In a similar way we may calculate the expected growth of this state due to additional neurons triggered by intracortical connections. If there were no subthreshold inputs from the sensory neurons, the fractional number of active neurons would be expected to increase by

$$\langle \alpha - \alpha_o \rangle \approx [1 - \exp(-\alpha_o \bar{\mu}_+) \sum_{l=0}^{\eta-1} (\alpha_o \bar{\mu}_+)^l / l!](1 - \alpha_0) \tag{3}$$

where α refers to the ACS, $\alpha = {}_a/A$, and η is the minimum number of average intracortical inputs causing a neuron to fire: $\eta = u(t/k_+)$. Equation (3) is only approximately correct because it has neglected the effect of the inhibitory connections and the effect of subthreshold inputs from the sensory neurons. If we take into account the first of these effects, equation 3 becomes

$$\langle \alpha - \alpha_o \rangle \approx (1 - \alpha_o) \exp(-\alpha_o h \bar{\mu}_-) \sum_{m=0}^{M} (\alpha_o h \bar{\mu}_-)^m / m! \cdot \{1 - \exp[\alpha_o(1 - h)\bar{\mu}_+] \sum_{l=0}^{\eta'-1} [\alpha_o(1 - h)\bar{\mu}_+]^l / l!\} \tag{4}$$

where $\eta' = u\left(\frac{t + mk_-}{k_+}\right)$. In the first summation it will be sufficient in general to consider only the first few terms, denoted here by M.

Finally, considering also the subthreshold inputs from the sensory neurons, we may write

$$\langle \alpha - \alpha_o \rangle = (1 - \alpha_o) \exp(-\sigma \bar{m} - \alpha_o h \bar{\mu}_-) \sum_{n=0}^{\xi-1} \sum_{m=0}^{M} (\sigma \bar{m})^n (\alpha_o h \bar{\mu}_-)^m / n! m! \cdot \{1 - \exp[-\alpha_o(1 - h)\bar{\mu}_+] \sum_{l=0}^{\eta''-1} [\alpha_o(1 - h)\bar{\mu}_+]^l / l!\} \tag{5}$$

where $\eta'' = u\left(\frac{t + mk_- - nc}{k_+}\right)$.

For the sake of illustration we show in Fig. 3 the results of equation 2 for the case, $\xi = 2$, i.e. the specific size of the FCS as a function of the specific size of the SS provided it takes at least two sensory-cortical inputs to trigger a cortical neuron.

The dependence of the specific size of the ACS on a given FCS (in the approximation of equation 3) is pictured in Fig. 4. We use the convenient parameters $\lambda = \alpha_o \bar{\mu}_+$ and $\gamma = \dfrac{\alpha - \alpha_o}{1 - \alpha_o}$. As we shall see, this set of curves is particularly useful in predicting the behavior of the network.

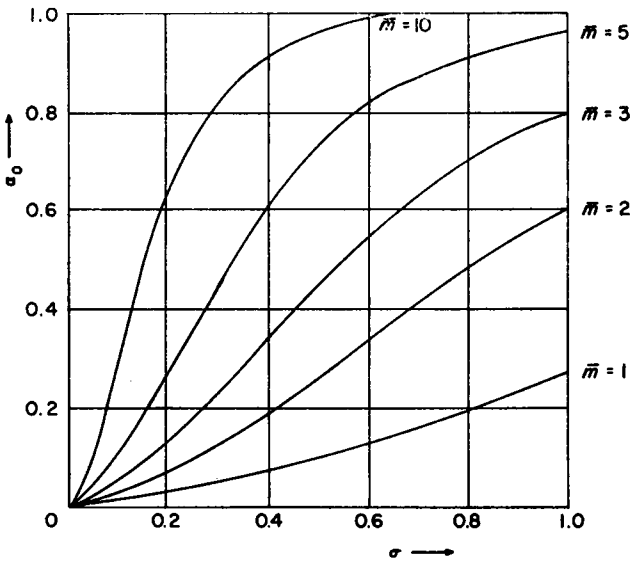

FIGURE 3. α_o, specific size of FCS, vs. σ, specific size of SS from equation 2, for $\xi = 2$ and different values of $\bar{m}$.

Reinforced Domains

The results of applying the reinforcement rules of equation (1) to a given state a will be to raise the intracortical coupling coefficients k_{ij} for a subset or domain of the A cortical neurons. Without loss of generality we can label the neurons in such a way that all neurons belonging to state a have consecutive numbers, say 1 through a. Similarly another, disjoint state may consist of neurons $(a + 1)$ through $(a + b)$, etc. The separate learning of these cortical states will affect the connectivity matrix in a way which is shown schematically in Fig. 5. Here the shaded areas represent reinforced domains, i.e. regions within which the nonvanishing matrix elements have been increased by δ, while in the blank regions the coefficients k_{ij} retain their previous (naive) values.

It should be pointed out that while this and many of the subsequent connectivity diagrams show symmetry with respect to the principal diagonal, this symmetry does not extend to the microstructure of the matrix which is fundamentally asymmetric because of the one-way character of the coupling coefficients.

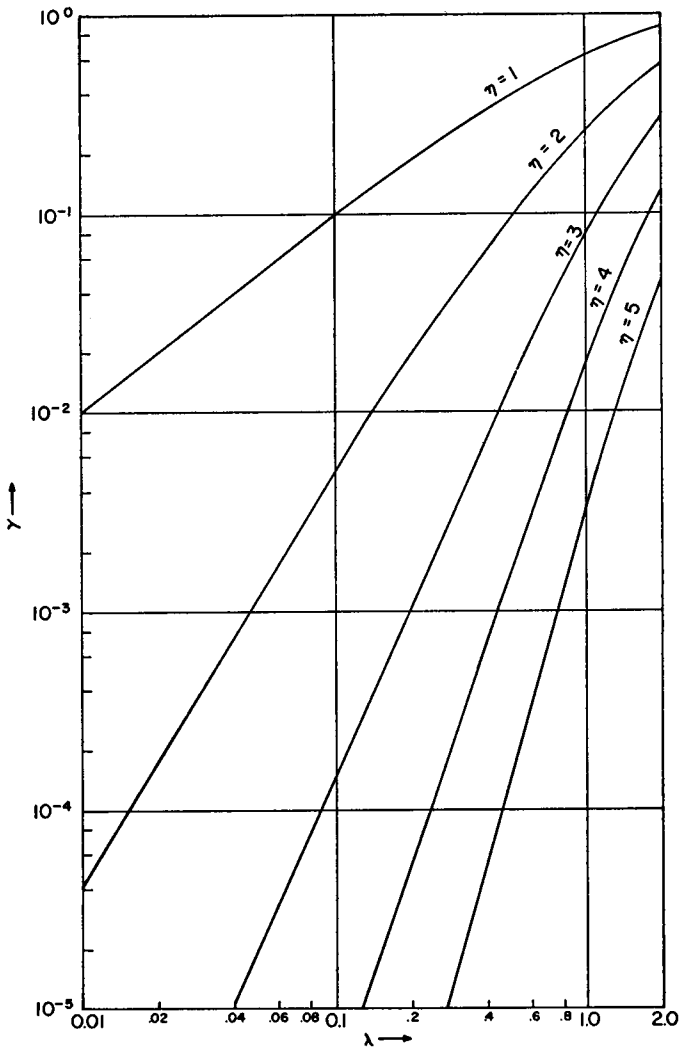

FIGURE 4. Curves of $\gamma = (\alpha - \alpha_o)/(1 - \alpha_o)$ vs. $\lambda = \alpha_o\bar{\mu}_+$ from equation 3. With the aid of these curves we obtain the specific sizes of ACS's for an FCS given by α_o. Alternatively, the same curves give γ vs. λ^* for a reinforced domain, as explained in the text.

The effect of reinforcing a domain a on the state resulting from an arbitrary stimulus may now be determined. Several cases are of interest.

Theorem 1. Reinforcing a state a, where $(a_o \mathsf{M} s_a)$ will leave the ACS resulting from s_a unchanged.

Proof: Because of the invariance of a_o we need only prove that the transition $a_o \rightarrow a$ remains unchanged. Let the new ACS be a'. But of all the connections coming out of the domain a_o only those are reinforced which lead to neurons in a. Since only

the excitatory connections are increased this causes additional positive inputs to neurons in a. Therefore any neuron belonging to a before the reinforcement, will belong to a' after reinforcement. On the other hand the inputs to any neuron not in a will be the same as before because of the identity of the active neurons acting as sources and the unchanged coupling coefficient. Hence any neuron not in a will remain inactive; therefore $a' = a$.

Theorem 2. If a state b is disjoint to state a ($a \cap b = 0$), then reinforcing a will *not* change the mapping $s_b \rightarrow b$.

Proof: Again b_o remains unchanged, and since b_o and a are disjoint none of the synapses originating in the domain b_o have been reinforced. Therefore $b' = b$.

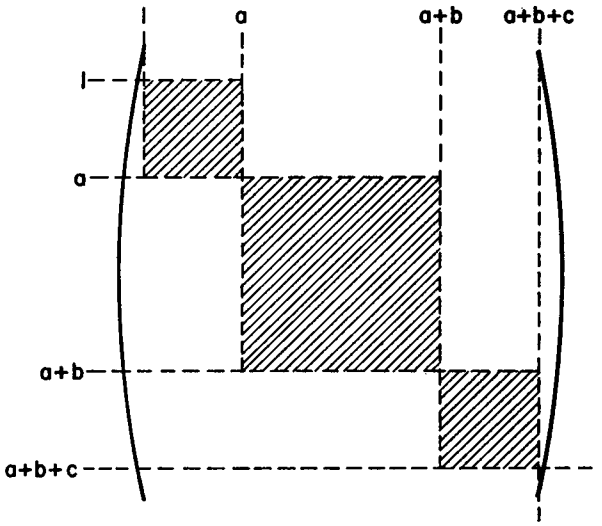

FIGURE 5. Schematic diagram of connectivity matrix $\{k_{ij}\}$, with reinforced domains shown as shaded areas.

Related States

We now consider the effect of the learning of a state a on a related state b, which however is not identical with a. In the Venn diagram (Fig. 6.) we show schematically the relation between a and b and the remaining cortical net A. The shaded region a is the reinforced region, i.e., all k_{ij} are increased if and only if $a_i \in a$ and $a_j \in a$. The switching on of the state b (e.g. by presenting the stimulus s_b) will now cause inputs into various domains in the net. Of these only the connections originating in $a \cap b$ and terminating in a (solid arrow in Fig. 6) are reinforced. All other connections are naive and will, on the assumption of a highly damped net, produce little significant activity. The effect of the reinforced connections leading from ($a \cap b$) to a may now be determined as follows: we shall neglect all nonreinforced inputs into the domain a. In this approximation we need only consider a net comprising just

the neurons in a and compute the spread of activity resulting when the subset $a \cap b$ is turned on. This is accomplished by rewriting equation 4 for the net a as follows:

$$\alpha^* - \alpha_o^* \approx (1 - \alpha_o^*) \exp(-\alpha_o^* h\mu_-^*) \sum_{m=0}^{M} (\alpha_o^* h\mu_-^*)^m/m!$$

$$\cdot\{1 - \exp[-\alpha_o^*(1-h)\mu_+^*] \sum_{l=0}^{\eta^*-1} [\alpha_o^*(1-h)\mu_+^*]^l/l!\} \qquad (6)$$

Here α_o^* refers to the ratio of the number of neurons in $a \cap b$ to that in a, α^* is the total fractional activity in a resulting from α_o^*; the quantities μ_+^*, μ_-^* and η^* for the

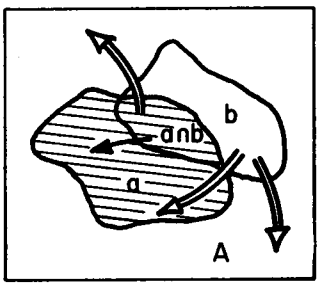

FIGURE 6. Venn diagram showing the effect of a reinforced domain a on a related state b. Efferent connections from b into various regions of the net are shown as arrows. The only connections that have been affected by the learning of b, are those leading from $a \cap b$ into a (solid arrow).

reinforced new a are obtained from parameters of the entire net by the relations

$$\mu_+^* = \frac{a}{A}\bar{\mu}_+; \qquad \mu_-^* = \frac{a}{A}\bar{\mu}_-; \qquad \eta^* = u\left(\frac{t + mk_-}{k_+ + \delta}\right) \qquad (7)$$

where a is the number of active neurons in a.

For $h = 0$ (no inhibitories) equation 6 reduces to a form equivalent to that of equation 3, and we may use the curves of Fig. 4 provided we use the starred quantities defined in equation 7. As an illustration consider a net of 10^5 neurons, a reinforced state a of 10^3 neurons and related state b of 10^3 neurons with an overlap $a \cap b$ containing 100 neurons. Let $\bar{\mu}_+ = 20$ and t and k_+ such that $\eta = 3$ for the naive net. In the reinforced domain a we let $\eta^* = 1$. The effect of turning on b may now be determined from Fig. 4. We consider first the activity caused throughout the net by the portion $\tilde{a} \cap b$ (where $\tilde{a}$ is the conjugate set of a). Since all connections originating in $\tilde{a} \cap b$ are naive we use $\eta = 3$ with $\alpha_o = 9 \times 10^{-3}$, $\lambda = 0.18$; from the graph (Fig. 4) we find $\gamma = 8 \times 10^{-4}$, hence about 80 additional neurons are expected to be triggered in the net; of these a fraction $900/10^5$ or less than one is expected to fall within a. In considering the effect of the reinforced connections we look at the net a. Here $\alpha_o^* = 0.1$, $\mu_+^* = 2.0$, and $\lambda^* = 0.2$. From the graph on the curve $\eta^* = 1$ we read $\gamma^* = 0.18$ and the additional number of neurons triggered in a becomes $1000 \times (0.9) \times (0.18) = 162$.

The example illustrates the very strong dependence of network behavior on η and shows that we were justified in disregarding all inputs into domain a not originating in $a \cap b$.

Preliminary Runs

We report here briefly the results of some computer simulation runs taken primarily to ascertain the relationships between sensory and cortical states, and the effect of inhibitory connections.

In Fig. 7 are plotted the specific sizes α as function of specific stimulus size σ. The data were obtained from a simulated net containing 100 sensory neurons and 900 cortical neurons. Other parameters were: $\bar{m} = 2$, $c = 20$, $\bar{\mu}_+ = 4$, $\bar{k}_+ = 10$, $\bar{\mu}_- = 6$, $\bar{k}_- = -10$, $h = 0.05$, $t = 40$. k_+ and k_- had flat distributions between 1 and 19, and -1 and -19 respectively. (a) α_o; (b) α; (c) fractional number of neurons having received subthreshold inputs; (d) fractional number of neurons affected by stimulus.

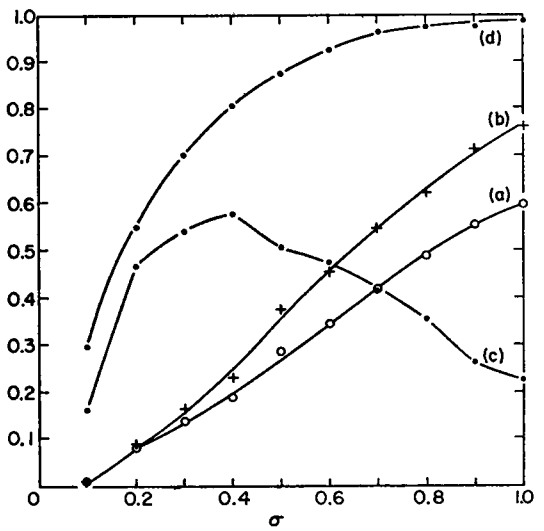

FIGURE 7. Specific size of cortical state (α) as function of specific size of sensory state (σ). The net parameters were $\bar{m} = 2, c = 20, \bar{\mu}_+ = 4, \bar{k}_+ = 10, \bar{\mu}_- = 6, \bar{k}_- = -10, h = 0.05, t = 40$. k_+ and k_- had flat distributions between 1 and 19, and -1 and -19 respectively. (a) α_o; (b) α; (c) fractional number of neurons having received subthreshold inputs; (d) fractional number of neurons affected by stimulus.

$\bar{k}_- = -10$, $h = 0.05$, $t = 40$. In this run the intracortical coupling coefficients k_+ and k_- were given flat distributions between 1 and 19, and between -1 and -19 respectively.

The curves show not only the sizes of the FCS's (α_o) and the ACS's (α) corresponding to different values of σ, but also the fractional number of cortical neurons that have received subthreshold inputs (curve c in Fig. 7) and the total fraction of the net involved, i.e., the sum of curves (b) and (c) in Fig. 7.

In Fig. 8 we show the effect of varying the fraction h of inhibitory neurons. For this run the network differed from the previous ones in that all excitatory intracortical coupling coefficients were equal to $+10$, all inhibitories to -10, and σ was constant at 0.4. The curves show again the specific sizes α_o and α corresponding to the FCS's and ACS's. The increased damping of the net with increasing h is apparent from the decrease of the quantity $\alpha - \alpha_o$.

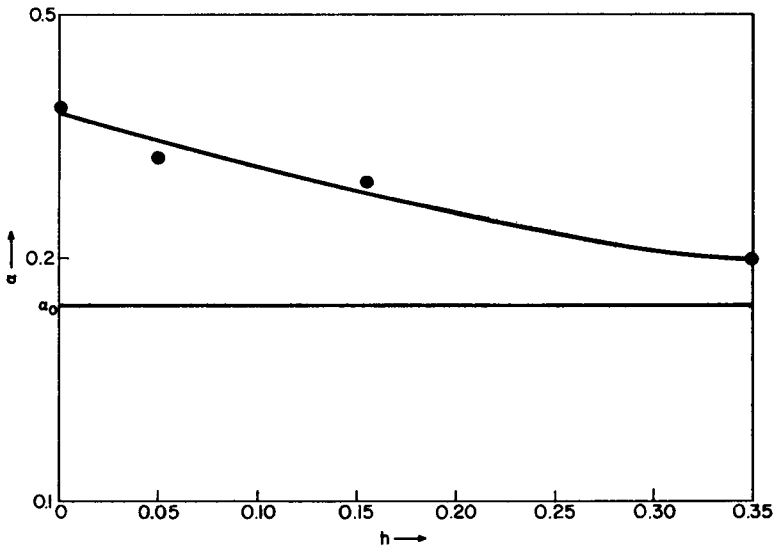

FIGURE 8. Specific size of cortical state α as function of inhibitory fraction h. The net used here is similar to the one in Fig. 7; σ was constant at 0.4.

ASSOCIATION OF CORTICAL STATES

In investigating the effect of learning we shall start with given FCS's, instead of generating them from the sensory states. We feel that the mapping between SS and FCS for our systems is sufficiently well understood to allow us to do this. This procedure will not only save us computing time, but allows us also to use larger cortical nets.

An FCS will thus be a selected set of cortical neurons, as though this set had been triggered by an SS. Our programs will enable us to generate for any FCS the corresponding ACS. Learning of such a state will consist of making the appropriate reinforcements as though synaptic facilitation occurred as a result of sustained activity in this ACS.

We now wish to define the association of two or more cortical states. Consider two stimuli s_a and s_b, and let their cortical representations in the naive net be the ACS's a and b respectively. Assume these states to be disjoint. We now form a new stimulus s_c which is the superposition of s_a and s_b, i.e., $s_c = s_a \cup s_b$. The FCS arising from this stimulus, call it c_o, clearly contains both a_o and b_o, hence the ACS c does also. If we now learn c, i.e. reinforce all excitatory connections k_{ij} within c, this will in general change the mapping of s_a and s_b into their respective ACS's. Call the new ACS's a' and b' respectively. From the considerations under Nomenclature and Calculations we conclude that under proper conditions a' and b' will be related to b_o and a_o respectively, i.e., $a' \mathbf{S} s_b$, $b' \mathbf{S} s_a$.

It is still true of course that $a' \mathbf{I} s_a$ and $b' \mathbf{I} s_b$, but from the considerations under

Nomenclature and Calculations we conclude that under proper conditions a' and b' will be related to b_o and a_o respectively, i.e., $a' \mathbf{S} s_b$, $b' \mathbf{S} s_a$. We call this operation the *binary association* of a and b. Similarly the association of more than two states involves the simultaneous showing of more than two stimuli and the learning of the resulting ACS.

The association between a and b can be represented schematically by showing the domain of reinforced matrix elements similar to Fig. 5. If a and b are initially disjoint states, we may again without loss of generality give all neurons in a and all neurons in b consecutive numbers so that all coupling coefficients belonging to one of the states form a submatrix in the connectivity matrix. The reinforced region resulting from the learning of c is shown in the schematic diagram in Fig. 9. Here the rows and columns are labeled according to the state to which they belong initially.

Two effects are not shown in the diagram but should be kept in mind. Because of the presence of inhibitory neurons the reinforced block shown in Fig. 9 may have some *holes* since there may be neurons belonging to a or b but not to c; also there

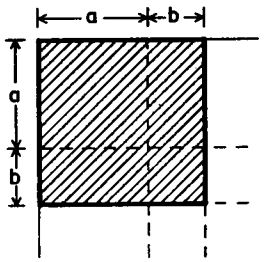

FIGURE 9. Schematic diagram of binary association of cortical states a and b. The shaded area represent that region in the matrix $\{k_{ij}\}$ within which all nonvanishing elements have been increased by δ.

will in general be a scatter of reinforced matrix elements outside the block $a \cup b$. Both effects will be small in the networks to be considered.

Simple Conditioned Reflex

We give here a qualitative argument showing how the classical conditioned reflex may arise in our model. More complex tasks will be studied in the next section by computer simulation of specific neural nets.

In the classical conditioned reflex a certain cortical state, call it the *unconditioned state* U always triggers a state R, which in turn leads to a definite observable motor, or glandular action. We take the neurons in R to be outside the random cortical association net, and assume that the wiring between U and R is to a large extent genetically determined. This has been called the *inborn reflex*. The situation is pictured in Fig. 10 by the asymmetrically located shaded area representing strong connections leading from U to R. Another state C, originally unrelated to either U or R, is now associated with U by presenting simultaneously the stimuli s_U and s_C. The resulting reinforcements are also shown in Fig. 10. As a result of the binary

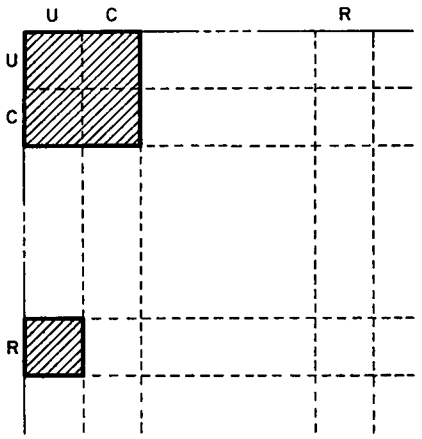

FIGURE 10. Schematic of simple conditioned reflex. Here U is the unconditioned state which has inborn connections to the reflex R (considered to be outside the random net of association neurons). A state C is connected to U by binary association.

association the conditioned stimulus s_c will have a new cortical response, call it c', where in general c' R U, or c' S s_u. It is clear that a sufficiently strong relationship c' R U may cause c' to activate enough of the strong native connections from U to R to elicit the observable response. The extent of the relationship c' R U may be determined by arguments similar to those given under Nomenclature and Calculations.

COMPUTER SIMULATION OF COGITIVE TASKS

A network of 1000 randomly interconnected neurons was generated by the computer. The fraction of inhibiting neurons was chosen to be 5%. All neurons had a threshold $t = 52$. The (naive) coupling coefficients were $k_+ = 10$, $k_- = -10$ and the average multiplicities of efferent connections were $\bar{\mu}_+ = 4$, $\bar{\mu}_- = 6$. The increment δ used in reinforcing excitatory synapses was 20. It can be seen from the curves in Fig. 4 (which are for $h = o$) that the parameters chosen make this net very highly damped, i.e., we find $\gamma \ll 1$ even for large initial excitation.

In each of the tasks we begin by selecting a number of FCS's which, for simplicity, have equal sizes and are disjoint. The neurons are labeled in such a way that each state is made up of consecutively numbered neurons. To illustrate a cognitive task we take each of these states to have a specific meaning. In each task the training involves a succession of binary or ternary associations whose effect on the connectivity matrix is shown in schematic diagrams. The performance of the tasks is seen in tabulations of the computer outputs.

In the first task we have selected 9 FCS's, each composed of 75 consecutive neurons and labeled A, A_1, A_2, B, B_1, B_2, C, C_1, C_2. Here A_1, and A_2 represent words in a language A, B_1 and B_2 words in a language B, C_1 and C_2 words in a language C. These facts are first taught to the network by the successive binary associa-

tions (A, A_1), (A, A_2), (B, B_1) etc. If A is the English language, then the learning of another language B is accomplished by associating each foreign word B_i with its English equivalent A_i. In the same way C is learned by the association pairs (C_1, A_1), (C_2, A_2). The resulting pattern of reinforced domains is shown schematically in Fig. 11.

The results of the computer simulation are shown in Table I. The first ten FCS's presented to the net, Ia and Ib in Table I, are training runs, i.e., the coupling coefficients for the resulting ACS's are reinforced; these are the binary associations. The table shows the sizes of the ACS's as well as the distribution of the activity among the various FCS's.

Following the training runs we show the performance of the trained net (IIa and IIb in Table I). As an example, the FCS $A_1 \cup B$ now grows into an ACS comprising 166 neurons. This state is relatively strongly related to the state B_1, and as shown in

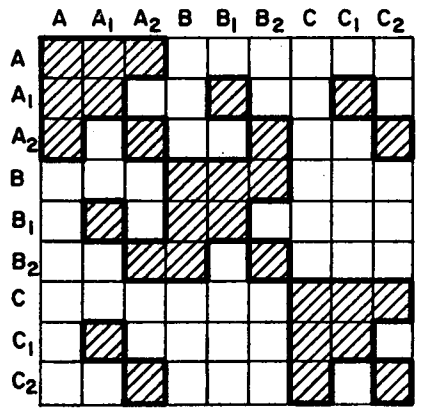

FIGURE 11. Association diagram resulting from language problem. Here A, B, C are languages, A_1 A_2, B_1 B_2, and C_1 C_2 are words in A, B, and C respectively. The diagram represents the learning of the foreign languages B and C.

the table, this relation exceeds that with other states. The ACS resulting from $A_1 \cup B$ thus contains the information[1] that B_1 is the equivalent to the English word A_1 in language B. The table shows that the correct answer is given every time, usually by a good margin.

In the last four examples, IIb in Table I, the overlaps are in general weak and spurious. This, of course, reflects the fact that the words in B have not been associated with their equivalents in C.

In our second task we use eleven disjoint states of 75 neurons each. The first three, labeled A, B, and C, represent objects that are to be classified according to properties

[1] We are not concerned in this paper with the neural mechanisms by which the information laid down in the ACS in the form of overlaps, is used to determine further neuronal activity. It is interesting, though, to make an analogy with quantum mechanics, regarding the ACS as a linear superposition of eigenstates with expansion coefficients which are simply the fractional overlaps of the ACS with the various orthogonal FCS's. Like the state function in quantum mechanics, the ACS is not an observable, but may be used to obtain observables.

X and Y. Each of the properties has three possible values: X_1, X_2, X_3, for X, and Y_1, Y_2, Y_3 for Y. State A is composed of the first 75 neurons of the net, B of the next 75 and so forth, leaving an unused residue of 175 neurons in the net. We assume that object A has the properties X_1 and Y_2, object B the properties X_1 and Y_3, and

TABLE I

SIMULATION OF LANGUAGE PROBLEM

In part Ia the net is taught the vocabulary of the three languages A, B, and C; in Ib the meaning of the foreign words B_1, B_2, C_1, C_2 is defined by binary associations with A_1 and A_2. The performance (without further learning) is shown in II, with the answerable questions in IIa. The correct answers are italicized; they coincide in all cases with the largest overlaps.

	Input FCS	Size of ACS	Distribution of activity								
			A	A_1	A_2	B	B_1	B_2	C	C_1	C_2
Ia	$A \cup A_1$	150	75	75							
	$A \cup A_2$	155	75	5	75						
	$B \cup B_1$	150				75	75				
	$B \cup B_2$	152				75	2	75			
	$C \cup C_1$	150							75	75	
	$C \cup C_2$	153							75	3	75
Ib	$A_1 \cup B_1$	157	3	75		4	75				
	$A_2 \cup B_2$	157			75	5	2	75			
	$A_1 \cup C_1$	160	3	75		1	6			75	
	$A_2 \cup C_2$	157			75			3	4		75
IIa	$A_1 \cup B$	166	3	75		75	*10*	2		1	
	$A_2 \cup B$	167			75	75	3	*10*			4
	$A_1 \cup C$	174	4	75		1	6		75	*7*	6
	$A_2 \cup C$	174			75			4	75	3	*17*
	$A \cup B_1$	166	75	*9*	2	4	75	1			
	$A \cup B_2$	176	75	6	*13*	5	2	75			
	$A \cup C_1$	166	75	*12*	1		3			75	
	$A \cup C_2$	175	75	6	*11*		1	1	5	1	75
IIb	$B_1 \cup C$	166	1	2		4	75		75	4	5
	$B_2 \cup C$	169			2	5	2	75	75	3	7
	$B \cup C_1$	163	1	6		75	4	2		75	
	$B \cup C_2$	165	1	1	6	75	1	2	4		75

object C the properties X_2 and Y_2. The network learns these facts through the sequence of binary associations (A X_1), (A Y_2), (B X_1), etc. It has previously been instructed that X_1, X_2, X_3 are values of the property X (e.g. red, green, and blue are colors) by the binary associations (X X_1), (X X_2) and (X X_3). Similarly for Y. The resulting association diagram is shown in Fig. 12.

The performance of the net following the 12 successive training instructions, is

shown in Table II. Presentation of the state $A \cup X$, for example, triggers an ACS which is most strongly related to X_1, while $A \cup Y$ points to Y_2, the correct Y-property of A. It is seen that similar inquiries elicit the appropriate answers for B and C. Conversely, the system is capable of naming the object (if such exists) corresponding to a given pair of properties. It should be noticed that, although the ACS's are fairly large, little or no activity spills over into the unoccupied portion of the net.

Conditioned Reflexes

A variety of conditioned reflexes have been reported following Pavlov's original work (Pavlov, 1960). We shall discuss here four types which were described by Asratian (1965). These are (*a*) the dual conditioned reflex, (*b*) the bidirectional

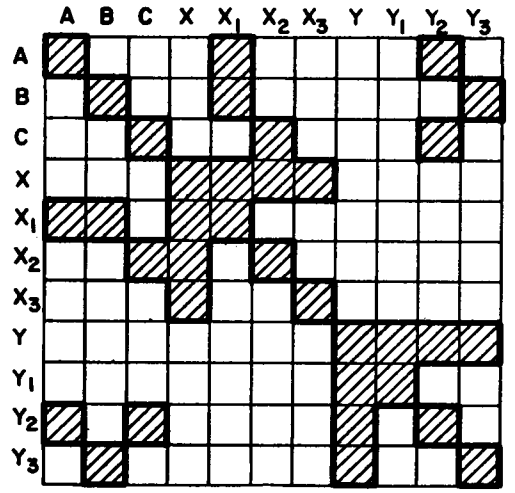

FIGURE 12. Association diagram resulting from the classification of three objects A, B, C according to properties X and Y. The net has learned that A has properties X_1, Y_2; B has X_1, Y_3; and C has X_2, Y_2.

conditioned reflex, (*c*) switching, and (*d*) the Struchkov reflex. We shall see that the first three are well within the range of capabilities of our model, while the fourth involves a degree of complexity that can only be achieved by the introduction of additional features. Only the switching reflex will be simulated.

Dual Conditioned Reflex. In this phenomenon two so-called unconditioned stimuli lead to different inborn reflexes. Associations are made alternately between conditioned stimulus and one or the other of the unconditioned stimuli until the latter by itself is able to elicit both of the reflexes. In the example given by Asratian the conditioned stimulus is the sound of a buzzer, the unconditioned stimuli are a mild electric shock and a food stimulus, leading, respectively, to salivation and withdrawal of the limb. If we let the cortical states corresponding to the stimuli be U_1 and U_2 (unconditioned states) and C, then, in the language of our model, the dual conditioned reflex may be described as a sequence of association pairs (U_1 C), and (U_2 C), leading to the association diagram in Fig. 13.

Bidirectional Conditioned Reflex. This phenomenon was described by Asratian as occurring when the conditioned stimulus itself has an inborn reflex. In the example given, a food stimulus is associated with a mild electric shock, thereby making the inborn reflex to the first become the conditioned reflex to the second. At the same time the reflex that is inborn for the second stimulus (here the withdrawal of a limb) will appear whenever the food stimulus is presented. This is called the *reverse conditioned reflex.* The situation is diagrammed in Fig. 14. It should be clear from the discussion that this and other reflexes described so far can be performed without difficulty by a network of the type used, for example, in connection with the language problem or object classification problem.

TABLE II

SIMULATION OF OBJECT CLASSIFICATION PROBLEM

In I objects A, B, C are associated with their properties through binary associations: A: X_1 Y_2 ; B: X_1 , Y_3 , C: X_2 , Y_2 . In II the performance of the net is shown. Italicized numbers are the correct answers.

	Input FCS	Size of ACS	Distribution of activity											Residue
			A	B	C	X	X_1	X_2	X_3	Y	Y_1	Y_2	Y_3	
	$X \cup X_1$	150				75	75							
	$X \cup X_2$	155				75	5	75						
	$X \cup X_3$	158				75	5	3	75					
	$Y \cup Y_1$	150								75	75			
	$Y \cup Y_2$	153								75	3	75		
	$Y \cup Y_3$	155								75	3	2	75	
I	$A \cup X_1$	155	75			4	75							1
	$A \cup Y_2$	154	75				1			2		75		1
	$B \cup X_1$	159	4	75		4	75							1
	$B \cup Y_3$	153	1	75			2						75	
	$C \cup X_2$	155			75	5		75						
	$C \cup Y_2$	162	2		75	1		5		3		75	1	
	$A \cup X$	182	75	3		75	*16*	2	3	1		6		1
	$A \cup Y$	183	75	2		1	3			75	3	*15*	8	1
	$B \cup X$	174	1	75	1	75	*14*	3	4				1	
	$B \cup Y$	168	1	75			2			75	3	3	*9*	
	$C \cup X$	177			75	75	9	*11*	4		2	1		
	$C \cup Y$	173		1	75	1		5		75	3	*7*	6	
II	$X_1 \cup Y_1$	168	4	3		4	75	1	1	4	75			1
	$X_1 \cup Y_2$	179	*13*	4	2	4	75	1		3		75	1	1
	$X_1 \cup Y_3$	168	4	*9*		4	75						75	1
	$X_2 \cup Y_1$	162			1	5		75		5	75		1	
	$X_2 \cup Y_2$	172	4		*8*	5		75		3		75	1	1
	$X_2 \cup Y_3$	161	1	4	1	5		75					75	
	$X_3 \cup Y_1$	160				5		1	75	4	75			
	$X_3 \cup Y_2$	169	4		3	6	1	2	75	3		75		
	$X_3 \cup Y_3$	161	1	3	1	4	1	1	75				75	

Switching. In this more complex situation we have two unconditioned stimuli, here denoted by s_{U1} and s_{U2}, with their resulting cortical states U_1 and U_2 and the inborn responses R_1 and R_2, as in the dual conditioned reflex. Again we have a neutral state, C, i.e., one initially disjoint with respect to U_1 and U_2 and incapable of eliciting either R_1 or R_2. In addition there are two other states[2] we shall call A_1 and A_2 which are also initially neutral. The training for the switching reflex consists of two successive triple associations (U_1A_1C) and (U_2A_2C). The effect on the k matrix is shown in the association diagram Fig. 15. Following the training it is observed that the animal will have the conditioned response R_1 if it is presented with

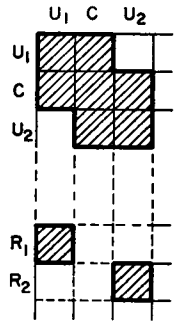

FIGURE 13. Diagram of dual conditioned reflex. U_1 and U_2 have inborn reflexes R_1 and R_2. Two binary associations are made: $(U_1\ C)$ and $(U_2\ C)$.

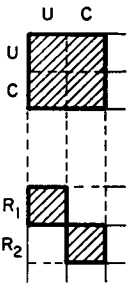

FIGURE 14. Diagram of bi-directional conditioned reflex. U and C have inborn reflexes R_1 and R_2 respectively. Association is made between U and C.

the stimulus $(s_C \cup s_{A1})$, and R_2 when receiving $(s_C \cup s_{A2})$. The states A_1 and A_2 thus have the effect of switching the conditioned response to C between R_1 and R_2.

The effect is readily understood in terms of our model. The triggering of a state $(C \cup A_1)$ will cause additional activity in both the reinforced domains $(U_1 \cup C \cup A_1)$ and $(U_2 \cup C \cup A_2)$, but the first is expected to be much stronger than the second because of the strong dependence of γ on λ shown in Fig. 4.

A computer simulation of the switching problem has produced the results shown in Table III. The network parameters are again the same as those used in the previous simulations. For the individual FCS's we have chosen disjoint states of 100

[2] In the switching experiments described by Asratian the states A_1, A_2 are some changing feature in the environment of the experimental animal, such as the identity of the experimenter, the time of day, or the box in which the animal is kept.

neurons each. The first two lines in Table III show the triple associations that constitute the learning phase, after which the switching is clearly demonstrated. Thus the ACS resulting from the input ($C \cup A_1$) is most strongly related to U_1, while ($C \cup A_2$) leads to an even stronger overlap with U_2. The approximate equation (6), incidentally, would predict an average overlap of 22 with the correct state and 7 with the *wrong* state.

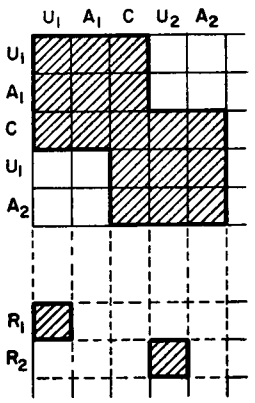

FIGURE 15. Diagram of switching. U_1 and U_2 have inborn reflexes R_1 and R_2 respectively. Two triple associations are made: (U_1 A_1 C) and (U_2 A_2 C). Here A_1, A_2, U_1, U_2 and C are originally disjoint states.

TABLE III

SIMULATION OF SWITCHING PROBLEM

I lists the triple associations which constitute the learning phase. In II the performance of the net is shown.

	Input FCS	Size of ACS	Distribution of activity				
			U_1	A_1	C	U_2	A_2
I	$U_1 \cup C \cup A_1$	300	100	100	100		
	$U_2 \cup C \cup A_2$	318	6	12	100	100	100
II	$C \cup A_1$	229	*12*	100	100	8	9
	$C \cup A_2$	243	7	12	100	*24*	100
	$C \cup U_1$	238	100	*20*	100	10	8
	$C \cup U_2$	239	7	12	100	100	*20*

Struchkov Reflex. This complex form of conditioned reflex was first reported by Struchkov (1956) and is described by Asratian (1965). It involves two unconditioned stimuli U_1 and U_2 (food and electroshock), two neutral stimuli C_1 and C_2 (buzzer and skin stimulus), and two environmental stimuli A_1 and A_2 (two different test chambers called *chamber 1* and *chamber 2*). At the outset the animal is taught two conditioned reflexes in chamber 1 by associating C_1 with U_1 and C_2 with U_2. As soon as stable conditioned reflexes were obtained the animal was placed in chamber 2 where the training continues, this time associating C_1 with U_2 and C_2 with

U_1. This extremely difficult situation is not handled too readily by the experimental animals, which tend to go through a neurotic state and a phase of double reflexes and wrong reflexes. In time, however, they will develop the responses appropriate to the chamber in which they received the conditioned stimulus.

Again, in the language of our model, the problem is represented by the triple associations $(U_1C_1A_1)$, $(U_2C_2A_2)$, $(U_1C_2A_2)$, and $(U_2C_1A_2)$. The corresponding association diagram is shown in Fig. 16. It will be noticed that the only nonreinforced areas connect the pairs (U_1U_2), (C_1C_2), and (A_1A_2). If we were to subject our net-

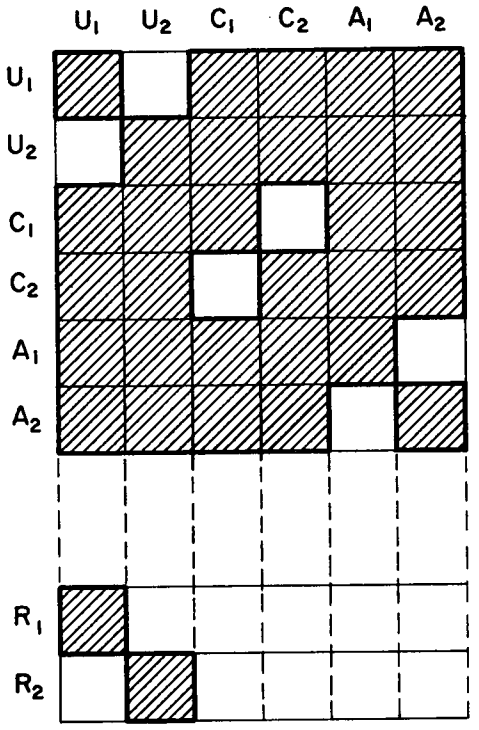

FIGURE 16. Diagram of Struchkov reflex showing the native connections to R_1 and R_2 and the following triple associations: $(U_1\ C_1\ A_1)$, $(U_2\ C_2\ A_2)$, $(U_1\ C_2\ A_2)$, and $(U_2\ C_1\ A_2)$.

work to the conditions of this experiment, we would find that it is completely and permanently incapable of distinguishing the right from the wrong response. Thus the input $(C_1 \cup A_1)$ will give the same expectation value for activity in domains U_1 as in U_2. It appears therefore that tasks of the degree of complexity of the Struchkov reflex cannot be accomplished by our model without added complexity.

LIMITATIONS AND EXTENSIONS

In concluding the discussion it is well to emphasize the limitations of the present model as well as the possibilities of extending it to include other features.

Perhaps the most obvious limitation is that of size, set at present by the limited

computer memory. However, in our treatment we were able to express the behavior of the net to a large extent in terms of intensive quantities, so that the scaling to larger nets is readily accomplished. What plagues experiments with the small nets is not so much that they behave fundamentally differently from their larger counterparts, but that their performance is more erratic owing to the greater importance of statistical fluctuations.

The assumption of a net that is homogeneous and isotropic and whose connections are completely random, is most certainly a gross oversimplification. If synaptic connections were to favor neighboring neurons, the response of the net to stimulation of a local group of neurons would be quite different from that of the random net. The effect is estimated in the following crude illustration: assume that in a local net the outputs from a small group of neighboring neurons effectively terminate within a local domain, comprising a fraction l of the total net of A neurons. The

TABLE IV

SPREAD OF ACTIVITY IN RANDOM AND LOCAL NET

Approximate values are obtained from Fig. 4 for a net of 10^{10} neurons in which 10^5 neurons are simultaneously triggered. For the local net it is assumed that synapses effectively go into a fraction $l = 10^{-2}$ of the total net. Multiplicities are $\mu_+ = 10^3$, $\mu_- = 0$.

	Spread of activity	
	Random net	Local net
$\eta = 1$	10^8	6.4×10^7
$\eta = 2$	4×10^8	2.7×10^7
$\eta = 3$	1.7×10^9	8×10^6

spread of activity within this domain may be estimated by using again the curves of Fig. 4, taking into account the effective values of μ_+ and μ_-. As an example we choose a net of 10^{10} neurons in which a local group of 10^5 neurons is turned on. The number of additional neurons triggered is given in Table IV for different values of the parameter η and under assumption (*a*) that the connections are random, and nonlocal and (*b*) that the connections are predominantly local, with the parameter l taken to be 10^{-2}. It is noted that with $\eta = 1$ the random net shows a greater spread of activity than the local net, because the latter is close to saturation. For higher values of η the very pronounced enhancement of activity in the locally connected net is evident. Other net parameters were $\mu_+ = 10^3$ and $\mu_- = 0$.

Considerations like the above should be useful in attempting to understand the effect of local cortical lesions as epileptogenic foci (Penfield and Rasmussen, 1950), or the spreading of locally induced activity across bridges of cortical material (Burns, 1958).

Time Dependencies

In our treatment we have completely neglected all temporal effects such as transmission and synaptic delays, refractory periods, afterdischarges and the like. This may seem too radical an assumption if one hopes to understand cortical activity, which is clearly continuous. It is recognized, however, that this continuity is due to a number of sources, such as the continuity of sensory input, coupling with subcortical systems, e.g. reverberatory thalamocortical circuits (Fair, 1963), spontaneous activity, burst response of some neurons to single triggers, and self-reexcitation in what Burns (1958) calls the type B network. Thus, to ascribe the observed continuous activity merely to an underdamped homogeneous net is not only a drastic oversimplification, but, moreover, one which should not be expected to yield much physical insight. Even forgetting about the effect of extracortical systems, the heterogeneity of the cortex makes it more realistic to think of it as a system of more or less distinct, but interwoven and interconnected nets having different properties and functions. Burns's type B network is one such example. Viewed in this way, the present study considers what may happen in a highly simplified version of one component of the system. This component is, as was repeatedly stressed, highly damped, hence could support sustained activity only through its interaction with other components. What we have demonstrated is the fact that a considerable amount of information processing can be accomplished in this component merely in the process of immediate mapping between what we called the sensory and cortical states, or, more specifically, the transformation that leads from the FCS to the ACS.

Memory Capacity

This problem is intimately connected with the so-called overwrite problem, the question to what extent additional learning erases previously acquired memories. In a sense the learning diagram of the Struchkov reflex (Fig. 16) is a picture of a sadly overloaded memory bank. The situation can be relieved in two ways: (*a*) by cortical inhibition and (*b*) by the fading of old memory traces. Various schemes of incorporating the first into our model are under consideration. The second involves questions regarding the time dependence of the physical changes that constitute the memory trace, a subject that was recently discussed by one of us (Harth, 1966b).

Spontaneous Activity

Burns defines this as the capability of a neuron to discharge "at moments which are not dictated by the precise times at which electrical energy is supplied from outside the system." Into this category belong the already mentioned afterdischarges and what we may call true spontaneous activity, the latter being either random or cyclic, but in any event unrelated to previous excitatory input. According to the picture presented by Burns, the afterdischarges occur in the critically damped net of type B

neurons; this component, once excited, may sustain activity for minutes. True spontaneous activity was reported to have been found in central neurons by Kristiansen and Courtois (1949). Random discharges may in principle be explained as fluctuations in the apparently random release of acetylcholine which causes the miniature end plate potentials observed in the case of motor neurons (Castillo and Katz, 1954), but some investigators doubt the existence of such activity in the cortex (Burns, 1958).

If truly spontaneous activity exists, it may be considered to feed into the cortical net a steady background of noise, which in the naive and highly damped net would produce little or no further activity. In the trained net, however, we may expect amplification of activity within the strongly reinforced domains. Such an effect could play a very important role in thinking processes.

The authors are indebted to the Computing Center of Syracuse University for time generously provided on the IBM 7074 Digital Computer.
In its later phase the research reported in this paper was sponsored by the Aerospace Medical Research Laboratories, Wright-Patterson AFB, under contract No. F33615-67-C-1413.

Received for publication 24 May 1967.

REFERENCES

ASHBY, W. R., H. VON FOERSTER, and C. C. WALKER. 1962. *Nature* **196**:561.
ASRATIAN, E. A. 1965. Compensatory Adaptations, Reflex Activity and the Brain. English transl. by S. A. Corson. Pergamon Press Ltd., Oxford, England.
BEURLE, R. L. 1956. *Phil. Trans. Roy. Soc. London*, Ser. B. **240**:55.
BIRKHOFF, G., and S. MACLANE. 1953. A Survey of Modern Algebra. The Macmillan Company, New York.
BURNS, B. D. 1958. The Mammalian Cerebral Cortex. Edward Arnold Ltd., London, England.
CAIANIELLO, E. 1961. *J. Theoret. Biol.* **1**:204.
DELCASTILLO, J., and B. KATZ. 1954. *J. Physiol.*, **124**:560.
ECCLES, J. C. 1953. The Neurophysiological Basis of Mind. The Clarendon Press, Oxford, England.
ECCLES, J. C. 1964. The Physiology of Synapses. Academic Press Inc., New York.
FAIR, C. M. 1963. The Foundations of the Psyche. Wesleyan University Press, Middletown, Conn.
FARLEY, B. G., and W. A. CLARK. 1961. Activity in networks of neuron-like elements. *In* Information Theory (Fourth London Symposium). C. Cherry, editor. Butterworth & Co. Ltd., London, England.
GRIFFITH, J. S. 1963. *Biophys. J.* **3**:299.
HARMON, L. D., and E. R. LEWIS. 1966. *Physiol. Rev.* **46**:513.
HARTH, E. M. 1966 *a*. Brain Models and Thought Processes. *In* Automata Theory. E. Caianiello, editor. Academic Press Inc., New York.
HARTH, E. M. 1966 *b*. On the nature and time dependence of memory. Second Intern. Biophysics Congress. Vienna, Austria. In press.
HEBB, D. O. 1949. The Organization of Behavior. John Wiley & Sons, Inc., New York.
HODGKIN, A. L., and A. F. HUXLEY. 1952. *J. Physiol.*, (London). **117**:500.
KRISTIANSEN, K., and G. COURTOIS. 1949. *Electroencephalog. Clin. Neurophysiol.* **1**:265.
LASHLEY, K. S. 1950. *Symp. Soc. Exp. Biol.* **4**:454.
LEITH, E. N., and J. UPATNIECKS. 1964. *J. Opt. Soc. Am.* **54**:1295.
LORENTE DE NÓ, R. 1949. The Structure of the Cerebral Cortex. *In* Physiology of the Nervous System. J. F. Fulton, editor. Oxford University Press, Inc., New York. 3rd edition.

McCULLOCH, W. S., and W. PITTS. 1943. *Bull Math. Biophys.* **5**:115.
VON NEUMANN, J. 1958. The Computer and the Brain. Yale University Press, New Haven, Conn.
PAVLOV, I. P. 1960. Conditioned Reflexes. G. V. Anrep, editor. Dover Publications, Inc., New York.
PENFIELD, W., and T. RASMUSSEN. 1950. The Cerebral Cortex of Man. The Macmillan Company, New York.
RAMÓN Y CAJAL, S. 1909. Histolgie du systemé nerveux de l'homme et des vertébrés. Maloine, Paris, France. **1**.
ROCHESTER, N., J. H. HOLLAND, L. H. HAIBT, and W. L. DUDA. 1956. *IRE Trans. Inform. Theory*. **2**: 80.
ROSENBLATT, F. 1962. Principles of Neurodynamics. Spartan Books Inc., Washington, D. C.
SMITH, D. R., and C. H. DAVIDSON. 1962. *J. Assoc. Comput. Mach.* **9**:268.
STRUCHKOV, M. I. 1965. *Pavlov J. Higher Nervous Activity* **6**:277, 282, 830 (1956).
WHITE, H. 1961. *Bull. Math. Biophys.* **23**:43.
ZEEMAN, E. C. 1962. The topology of the brain and visual perception. *In* Topology of 3 - Manifolds. M. K. Fort Jr., editor. Prentice-Hall Inc., Englewood Cliffs, N. J.